

November 1996
 FSU-SCRI-96-134
 IASSNS-HEP 96/115

Improving Lattice Quark Actions

M. Alford

*School of Natural Sciences
 Institute for Advanced Study
 Princeton, NJ 08540*

T.R. Klassen

*SCRI
 Florida State University
 Tallahassee, FL 32306-4052*

G.P. Lepage

*Newman Laboratory of Nuclear Studies
 Cornell University
 Ithaca, NY 14853*

Abstract

We explore the first stage of the Symanzik improvement program for lattice Dirac fermions, namely the construction of doubler-free, highly improved classical actions on isotropic as well as anisotropic lattices (where the temporal lattice spacing, a_t , is smaller than the spatial one). Using field transformations to eliminate doublers, we derive the previously presented isotropic D234 action with $\mathcal{O}(a^3)$ errors, as well as anisotropic D234 actions with $\mathcal{O}(a^4)$ or $\mathcal{O}(a_t^3, a^4)$ errors. Besides allowing the simulation of heavy quarks within a relativistic framework, anisotropic lattices alleviate potential problems due to unphysical branches of the quark dispersion relation (which are generic to improved actions), facilitate studies of lattice thermodynamics, and allow accurate mass determinations for particles with bad signal/noise properties, like glueballs and P-state mesons. We also show how field transformations can be used to completely eliminate unphysical branches of the dispersion relation. Finally, we briefly discuss future steps in the improvement program.

1 Introduction

Lattice calculations suffer from scaling errors, or lattice artifacts, that typically decrease like some power, a^n , of the lattice spacing (ignoring logarithmic corrections). Continuum results are obtained as $a \rightarrow 0$, but the cost of a realistic simulation of QCD, for example, grows like some large power, $a^{-\omega}$, of the inverse lattice spacing (ω is at least 6, but could even be 10 or more [1]). It is therefore extremely important to find highly improved actions; they will give accurate results on much coarser lattices.

Classical field theory estimates suggest that eliminating errors through order a^2 and maybe a^3 allows one to model the properties of a smooth bump with errors of a few percent to a fraction of a percent by using a lattice with 3–6 grid points per diameter of the object in each direction. For hadrons this means that spatial lattices of spacing $a = 0.2 - 0.4$ fm might suffice for improved actions (whereas $a = 0.05 - 0.1$ fm are typically used for unimproved actions). Even considering the computational overhead due to the more complicated form of improved actions, it is clear that being able to work on coarse lattices would save many orders of magnitude in CPU time.

For pure glue it has already been demonstrated [2] that this is possible. In this paper we describe some of the steps that are necessary to extend these savings to the more difficult problem of lattice quark actions. Our considerations are mainly classical, but we will outline where quantum effects seem to play a role and how to take them into account. The approach we would like to follow, pioneered by Symanzik [3, 4, 5, 6, 7], is to try to improve lattice actions and fields to some finite order in a , like a^2 or a^4 .

For asymptotically free theories, such as QCD, the terms in the action can be organized by their UV dimensions. Symanzik improvement then consists of adding higher dimension improvement terms to the action, mimicking the effects of the UV modes left out on the lattice. To fix the coefficients of these terms, one then proceeds as follows. Write down all terms with the appropriate symmetries up to the desired order, with arbitrary coefficients. Tune the coefficients by matching to a sufficiently large set of observables, calculated either in perturbation theory or non-perturbatively in a Monte Carlo simulation. Once the tuning is completed, the coefficients in the action will be functions of the physical couplings and a set of redundant couplings (see below). Improving (composite) field operators involves a similar process, which must be repeated for each independent field.

The above program is quite difficult to carry through in practice, not only non-perturbatively, but even in perturbation theory. In the past, standard lattice perturbation theory suffered from the rather debilitating problem that it did not seem to work very well, at least compared to continuum perturbation theory. This has now largely been understood [8] as being due to large renormalizations from tadpole diagrams, which occur in (naive) lattice but not continuum perturbation theory. Using a simple mean-field type method, known as tadpole improvement, one can design more continuum-like operators in which the tadpole contribution is greatly reduced. Tadpole improvement has been shown to work well for a variety of actions on surprisingly coarse lattices [2, 9, 10, 11, 12, 13, 14, 15].

We emphasize that tree-level tadpole improvement should be thought of as a first step in a systematic procedure of improving lattice actions. The next step can be further perturbative improvement, or, if there are reasons to believe that this is not sufficient, non-perturbative improvement.

It turns out to be substantially harder to improve lattice fermions than gluons, even on the classical level. The reason, ultimately, is the first-order nature of the fermion field equations, which leads to the well-known doubler problem, which we will discuss later. For Dirac fermions (quarks), Wilson [16] solved this problem by adding a second-order derivative term to the action. This term breaks chiral symmetry at $\mathcal{O}(a)$. Such errors — which are much larger than the $\mathcal{O}(a^2)$ errors of “naive lattice fermions” — are too large for this action to be useful in coarse lattice simulations.

The point of this work is to present doubler-free quark actions, for light *and* heavy quarks, that are classically improved to high order. The general tool to construct such actions will be *field redefinitions*; they allow one to introduce second-order derivative terms without destroying improvement. To allow the simulation of heavy quarks — and also to avoid potential problems due to unphysical branches of the quark dispersion relation, which are generic to improved actions — we can use *anisotropic lattices* [17, 18]. Let us discuss these ideas in turn.

Field redefinitions are just changes of variable in the path integral, so they do not affect spectral quantities (at least if one takes into account the Jacobian). Off-shell quantities of course do change. Since field redefinitions involve one or more free parameters, they lead to so-called *redundant* couplings, whose values can be adjusted at will. This freedom can be used to solve the doubler problem, for example. In other situations, in particular on the quantum level, it is very convenient to simplify an improved action by setting certain couplings to zero. This leads to the concept of *on-shell improvement*, where only spectral quantities can be obtained directly from the action (by improving composite operators one can, however, also obtain their matrix elements between physical states).

The “canonical” procedure for obtaining a doubler-free quark action correct up to, say, $\mathcal{O}(a^n)$ classical errors involves the following three steps:¹

1. Start with the continuum Dirac action and apply a field redefinition introducing even-order derivative terms into the action.
2. Expand the continuum operators in the transformed action in terms of lattice operators up to $\mathcal{O}(a^n)$ errors; this step will be referred to as the *truncation*. The even-order lattice derivative terms will eliminate the doublers that would be present without the field redefinition. One can stop here if one is only interested in spectral quantities; they will be classically correct up to $\mathcal{O}(a^n)$ errors.
3. To classically also improve off-shell quantities, *undo* the field transformation (now on the level of the lattice action). The resulting action and fields differ only at

¹The simple recipe to follow is a streamlining of what can be found in [6] together with [19].

$\mathcal{O}(a^n)$ form their continuum counterparts, and, in contrast to a naive discretization, have no doublers.

We emphasize that the improved actions so constructed are (classically) improved in every respect; the improvement of interactions does not have to be checked separately. When applied to lowest order ($n = 2$) the above procedure gives the Sheikholeslami-Wohlert action, originally suggested [6] as an improvement of the Wilson action. The next order ($n = 4$) yields the class of “D234” actions (in addition to the second order derivative Wilson term, they also contain third and fourth order derivative terms).

As alluded to earlier, a problem generic to actions improved beyond $\mathcal{O}(a)$ is the existence of unphysical branches of the free dispersion relation, simply due to higher order time derivatives in the action. We will refer to these extra branches as lattice *ghosts*. Their energies are at the scale of the (temporal) cutoff, so they will eventually decouple as the lattice spacing is decreased. For the lattice spacings used in practice their effect on, say, the hadron spectrum has not been thoroughly studied, but they certainly affect the renormalization of the improvement terms in the action. In addition, they can complicate numerical simulations by introducing oscillations in correlation functions at small times.

If either of these issues turns out to be a problem, one can deal with the ghosts in one of two ways. Firstly, one can use field transformations to replace the temporal with spatial derivatives. This produces somewhat more complicated actions, as we will see. Alternatively and secondly, one therefore might want to use anisotropic lattices with smaller temporal than spatial lattice spacing, $a_t < a_s$, to push up the energy of the ghosts and decouple them.

Besides effectively solving the potential problem of ghost branches, the use of anisotropic lattices has other advantages:

- By choosing a_t sufficiently small, one can simulate heavy quarks within a relativistic framework [11] without the prohibitive cost of a fine spatial lattice. This provides a simple alternative to the NRQCD [20] and Fermilab [21] formalisms.
- The signal to noise ratio of a correlation function calculated in a Monte Carlo simulation decays, generically, exponentially in time. Choosing a smaller a_t gives more time slices with an accurate signal, allowing for more precise and confident mass determinations. This is important for particles with bad signal/noise properties, like P-state mesons [22] and glueballs [23].
- It facilitates thermodynamic studies — one of the reasons being simply that it is easier to take independent derivatives with respect to volume and temperature if one can vary a_t independent of a_s — especially at high temperatures.
- It allows for significant simplifications in the design of improved actions. This will be relevant for our D234 actions.

All these advantages come at a price. Because they have lost part of their axis-permutation symmetry, anisotropic actions have more independent coefficients. This is not a problem at the classical level, but at the quantum level some of these coefficients will have to be tuned to restore space-time exchange symmetry. Large renormalizations violating space-time exchange symmetry were in fact seen in first attempts of using anisotropic lattices, see [17] and references therein. We find that with an improved gluon action and a tadpole improvement prescription appropriate to anisotropic lattices, such effects are relatively small, at the level of several percent on coarse lattices [18].

So far we have concentrated on classical Symanzik improvement; however, the improvement of fermion actions is also more difficult on the quantum level. For pure glue, the largest error at order a^2 is the violation of rotational invariance, which can be tuned to zero non-perturbatively, by demanding rotational invariance of the static potential at physical distances.² Actually, it seems that most of these errors are removed by tadpole improvement [2, 18].

Wilson-type quarks, on the other hand, have $\mathcal{O}(a)$ errors on the quantum level, and to eliminate them one has to tune a term that violates chiral but not rotational symmetry. The leading a^2 errors behave in the opposite way; they violate rotational symmetry but not chiral symmetry (so they are similar to the errors of gluons). The $\mathcal{O}(a)$ and (leading) a^2 errors of Wilson-type quarks therefore have very distinct effects on spectral quantities, and can be tuned iteratively, even on a non-perturbative level, by demanding chiral symmetry for the former, and rotational symmetry for the latter.

As for glue it seems that tadpole improvement does quite a good job in estimating the coefficient of the $\mathcal{O}(a^2)$ terms that lead to a restoration of rotational symmetry. Concerning the restoration of chiral symmetry to eliminate $\mathcal{O}(a)$ quantum errors, Lüscher et al have recently shown in some beautiful work [7] how to implement this in practice for the case of SW quarks on Wilson glue.

The outline of the remainder of this paper is as follows. In sect. 2 we discuss naive lattice fermions, doublers and ghosts. We proceed in sect. 3 to describe in more detail the three-step procedure to eliminate doublers, which we then apply to derive the Sheikholeslami-Wohlert and D234 actions on a general anisotropic lattice. Several special cases and variations are also discussed, including a completely ghost-free D234-like action. In sect. 4 we investigate the large mass behavior of the D234 actions. We conclude in sect. 5 with a brief summary and sketch of future steps in the improvement program.

Appendices A and B summarize our notation for euclidean continuum and lattice QCD, respectively. The reader might want to skim these appendices before starting with the main text, and later refer back to them as necessary. Appendix C discusses the dispersion relation of the D234 actions. Finally, in appendix D we give some formulas

²To on-shell improve gluons at order a^2 on an isotropic lattice one has to add two terms to the leading plaquette, which one can choose to be the “rectangle” and the “parallelogram” [4]. It is a certain linear combination of the coefficients of these two terms that can be tuned non-perturbatively by demanding rotational invariance. Since the coefficient of the parallelogram seems to be very small (it certainly is to one loop) [5, 2], this amounts to an “almost full” non-perturbative tuning of the glue at order a^2 .

useful in the tadpole improvement of the D234 actions.

Brief accounts of various parts of this work have appeared earlier in [10, 11, 24, 25].

2 Naive Lattice Fermions, Doublers, and Ghosts

Discussions of Dirac fermions on the lattice³ usually start with the so-called *naive* lattice fermions, specified by the fermion operator $\nabla + m$. Here ∇_μ is the usual first order, anti-hermitean, covariant lattice derivative,

$$\nabla_\mu \psi(x) \equiv \frac{1}{2a_\mu} \left[U_\mu(x) \psi(x + \mu) - U_{-\mu}(x) \psi(x - \mu) \right] \quad (2.1)$$

in terms of the link field $U_\mu(x)$ (cf. appendix B for details). This derivative differs at order a_μ^2 from the continuum Dirac operator $D_\mu = \partial_\mu - iA_\mu$.

One way of stating the origin of the doubler problem is that ∇_μ *decouples even and odd* sites of the lattice. This leads to a doubling (per direction) of the number of degrees of freedom on the lattice. If it were not for this problem, naive fermions would provide a lattice discretization of Dirac fermions with order a^2 errors. Similarly, one could use an improved operator, such as

$$\nabla_{c\mu} \equiv \nabla_\mu \left(1 - \frac{1}{6} a_\mu^2 \Delta_\mu \right) = D_\mu + \mathcal{O}(a_\mu^4), \quad (2.2)$$

where Δ_μ is the standard second order lattice derivative of appendix B, and the subscript “c” stands for “continuum-like”. The fermion operator $\nabla_c + m$ defines what we will refer to as *naive improved* lattice fermions. They would provide a lattice action with only order a^4 errors — again, if we could ignore the doublers.

The simplest way of discussing the doubler problem for a generic lattice action is in momentum space. Let us consider the dispersion relation $E = E(\mathbf{p})$, obtained from the poles of the free euclidean propagator, $p_0 = p_0(\mathbf{p})$, via $E = \pm i p_0$. The two signs correspond to particle and anti-particle. For simplicity we will factor out the sign and consider only solutions where the (real part of the) energy is positive. The quantitative details of the dispersion relations of the actions considered in this paper are discussed in appendix C. In figure 1 we show the massless dispersion relations of naive and naive improved fermions on an isotropic lattice. There are several noteworthy features.

- For naive fermions the one branch of the dispersion relation presented in figure 1 is purely real. Since we can only exhibit a cross section of the energy surface, one sees only one of the *spatial* doublers, which account for half of the doublers. The term “spatial doubler” refers to the fact that for each possible energy E there are generically eight momenta \mathbf{p} (with all components positive) such that $E = E(\mathbf{p})$.

³It is even more difficult to put *chiral* fermions on the lattice [26]. This we will not attempt. See [27] for recent work on this problem.

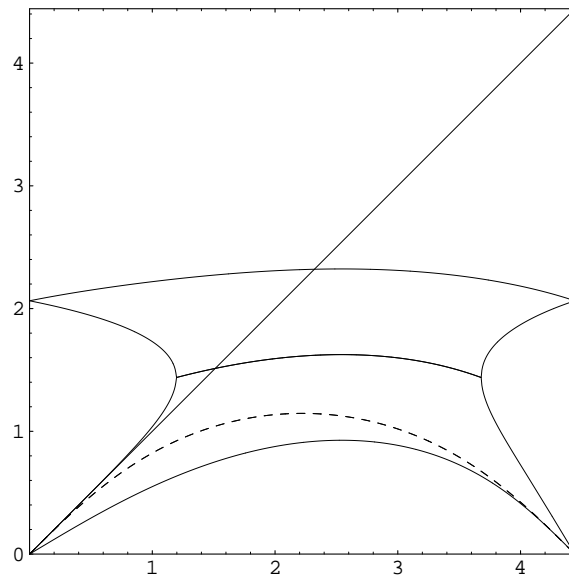


Figure 1: (Real part of the) massless dispersion relation $aE = aE(\mathbf{p})$ as a function of $a|\mathbf{p}|$, with $\mathbf{p} \propto (1, 1, 0)$, for naive fermions (dashed) and naive improved fermions (solid) on an isotropic lattice. For comparison we also show the dispersion relation of continuum fermions (thin solid).

- We also can not show that for each possible $E = E(\mathbf{p})$ of naive fermions there is another pole of the propagator at $E + i\pi/a_0$. This (as well as the existence of the spatial doublers) follows from the fact that in momentum space the action of naive fermions only depends on $a_\mu \bar{p}_\mu \equiv \sin(a_\mu p_\mu)$, which is invariant under $a_\mu p_\mu \rightarrow \pi - a_\mu p_\mu$. These complex poles constitute the *temporal* doublers.
- For naive improved fermions the picture is more complex. There are now four branches. The lowest branch is somewhat pathological in that its imaginary part is π/a for all momenta and in that its real part is *lower* than that of the physical branch. It is easy to see that this branch is related to the temporal doubler of the naive fermion action.

Clearly, neither of these two actions corresponds to what one might expect a lattice Dirac fermion to look like. Due to the doublers, the naive fermion action actually describes 16 Dirac fermion species in the continuum limit [28]. In addition to spatial doublers, naive improved fermions have *ghosts* (or *lattice ghosts*, if confusion could arise), as extra branches of the dispersion relation will be called from now on.⁴ As mentioned

⁴To be sure, these ghost branches are not related to the ghosts appearing in loop diagrams of perturbation theory in non-abelian gauge theories. These branches do not describe independent particles; they are just lattice artifacts related to the lattice “particle” described by the physical branch of the

in the introduction, if the ghosts should turn out to be a problem, there are always ways of eliminating them, to be discussed in the next section.

A solution to the doubler problem was proposed by Wilson [16], who suggested to add a second-order derivative term (now known as the Wilson term) to give the fermion operator

$$M_W = m_0 + \sum_{\mu} \left(\gamma_{\mu} \nabla_{\mu} - \frac{1}{2} r a \Delta_{\mu} \right), \quad (2.3)$$

on an isotropic lattice. r is the so-called Wilson parameter. It is easily shown (cf. sect. 3) that whereas there is one ghost branch for generic r , there is none for $r=1$. There is no doubler problem for any $r > 0$.

The Wilson action with $r = 1$ therefore solves the doubler problem without introducing any ghosts. However, the addition of the Wilson term introduces $\mathcal{O}(a)$ errors, which are too large for this action to be useful on coarse lattices. Sheikholeslami and Wohlert [6] described a modification of the Wilson action that has only $\mathcal{O}(a^2)$ errors for on-shell quantities. Their action differs from the Wilson action by a $\sigma F \equiv \sum_{\mu\nu} \sigma_{\mu\nu} F_{\mu\nu}$ term (cf. appendices A and B for notation), commonly known as the *clover* term. The free dispersion relation of this action is therefore identical to that of the Wilson action.

In [6] only on-shell improvement was considered. Later it was realized [19] that by performing a suitable change of variables on the fields in the Sheikholeslami-Wohlert action, one can also calculate off-shell quantities up to $\mathcal{O}(a^2)$ errors, at tree level. We will present a succinct derivation of all this, and its generalization to higher orders of improvement on anisotropic lattices, in the next section.

3 Improved Lattice Fermion Actions

3.1 Improvement without Doublers

As seen in the previous section, the naive and naive improved fermion actions have a doubler problem. More generally, this is true for any fermion matrix of the form

$$\sum_{\mu} \gamma_{\mu} \nabla_{\mu} (1 - b_{\mu} a_{\mu}^2 \Delta_{\mu} + d_{\mu} a_{\mu}^4 \Delta_{\mu}^2 + \dots), \quad (3.1)$$

which preserves chiral symmetry. (And even more generally, this follows from the Nielsen-Ninomiya theorem [26]; see also [28].) To avoid doublers we therefore should, following Wilson, introduce chiral symmetry breaking, even-derivative terms Δ_{μ} (or powers thereof) into the action. However, we would like to avoid the $\mathcal{O}(a)$ errors that a naive addition of a Wilson term entails. This can be achieved by a *field redefinition*.

The simplest way to proceed is to perform the field redefinition in the continuum and

dispersion relation. What justifies naming them ghosts is that they usually (but not always) give negative contributions to the spectral representation of correlation functions. In practice this leads to a characteristic “dip” in effective mass plots.

only subsequently discretize the action. Starting with the continuum action

$$\int \bar{\psi}_c M_c \psi_c \equiv \int d^4x \bar{\psi}_c(x) (\not{D} + m_c) \psi_c(x) , \quad (3.2)$$

we perform a field redefinition

$$\begin{aligned} \psi_c &= \Omega_c \psi \\ \bar{\psi}_c &= \bar{\psi} \bar{\Omega}_c \\ \bar{\psi}_c M_c \psi_c &= \bar{\psi} M_\Omega \psi , \quad M_\Omega \equiv \bar{\Omega}_c M_c \Omega_c . \end{aligned} \quad (3.3)$$

Note that a field transformation does not affect spectral quantities, at least if we take into account the Jacobian of the transformation. Classically the Jacobian does not matter. On the quantum level its leading effect is to renormalize the gauge coupling.

Our canonical choice of field redefinition is (with $\bar{\Omega}_c$ acting to the right)

$$\bar{\Omega}_c = \Omega_c , \quad \bar{\Omega}_c \Omega_c = 1 - \frac{ra_0}{2} (\not{D} - m_c) . \quad (3.4)$$

At this point a_0 is just a constant with the dimension of length, but in the subsequent lattice discretization a_0 will become the temporal lattice spacing.

The transformed fermion operator M_Ω reads

$$\begin{aligned} M_\Omega &= m_c + \not{D} - \frac{1}{2} ra_0 (\not{D}^2 - m_c^2) \\ &= m_c(1 + \frac{1}{2} ra_0 m_c) + \not{D} - \frac{1}{2} ra_0 \left(\sum_\mu D_\mu^2 + \frac{1}{2} \sigma \cdot F \right) , \end{aligned} \quad (3.5)$$

where we used eq. (A.11). We can *now* put the above action on the lattice by discretizing \not{D} and D_μ^2 and $F_{\mu\nu}$ to some order, a^n , using, for example, eqs. (B.9), (B.10), (B.15) and (B.16) in the $n=4$ case. Let us call the lattice action so obtained M .

If one is only interested in spectral quantities, one can use the propagator $G = M^{-1}$ in further calculations. Off-shell quantities will then generically have $\mathcal{O}(a)$ errors, since $\Omega_c = 1 + \mathcal{O}(a)$ and therefore $\psi = \psi_c + \mathcal{O}(a)$. However, as our third step, we can also improve off-shell quantities by *undoing* the field transformation. To do so, we use the obvious lattice versions of the operators in eq. (3.4), which differ from them at order a^n . Let us call these operators Ω and $\bar{\Omega}$. The action obtained by undoing the change of variable is $\bar{\Omega}^{-1} M \Omega^{-1}$, using fields that differ from the original continuum fields only at $\mathcal{O}(a^n)$. The propagator of this action is

$$G = \Omega M^{-1} \bar{\Omega} = M_c^{-1} + \mathcal{O}(a^n) . \quad (3.6)$$

Note that undoing the field transformation on the lattice does not lead to the (re)appearance of doublers.

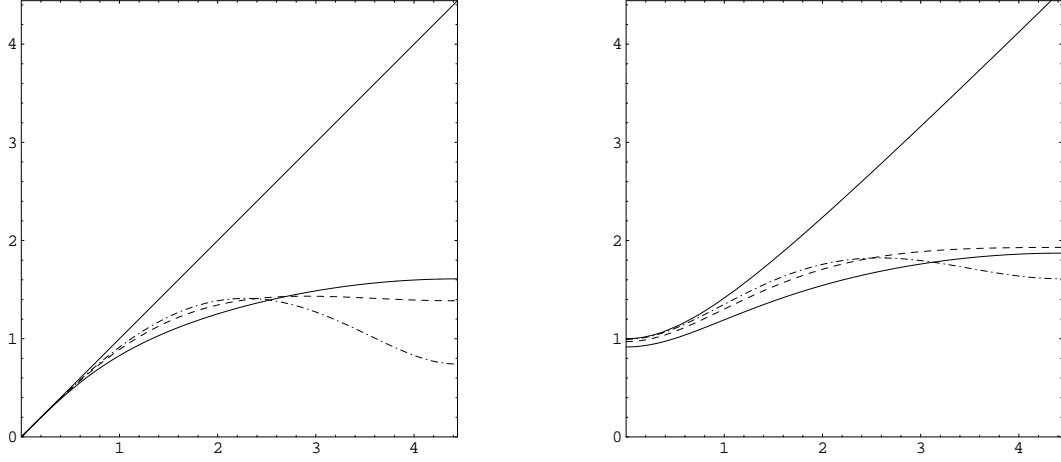


Figure 2: The energy $a_s E(\mathbf{p})$ of free SW/Wilson fermions with $r=1$ as a function of $a_s |\mathbf{p}|$ with $\mathbf{p} \propto (1, 1, 0)$. On the left we show the massless case on 1:1 (solid), 2:1 (dashed), and 5:1 (dot-dashed) lattices, as well as continuum fermions (thin solid). On the right we show the same for mass $a_s m_c = 1$.

3.2 The Sheikholeslami-Wohlert Action and $\mathcal{O}(a)$ Terms

Using the leading discretization of the derivatives in (3.5) gives

$$M_{\text{SW}} = m_c(1 + \frac{1}{2}ra_0m_c) + \not{V} - \frac{1}{2}ra_0 \sum_{\mu} \Delta_{\mu} - \frac{1}{4}ra_0 \sigma \cdot F. \quad (3.7)$$

This is the Sheikholeslami-Wohlert (SW) action on an anisotropic lattice. For $F_{\mu\nu}$ one can use the so-called *clover* representation (cf. appendix B), which has $\mathcal{O}(a^2)$ errors. By construction this action has classical $\mathcal{O}(a^2)$ errors for spectral quantities; also for off-shell quantities if we undo the change of variable. To obtain the Wilson action one must by hand set the clover term, $\sigma \cdot F$, to zero, thereby incurring an $\mathcal{O}(a_0)$ error in the presence of a non-trivial gauge field.

This action has no doublers for any $r > 0$; it has no ghost branches either if and only if $r=1$, which is therefore the canonical choice. For future reference we show in figure 2 the free dispersion relations for the SW/Wilson action with $r=1$ for various anisotropies a_s/a_t and masses.

Recall that in general more operators are needed for quantum improvement (even on-shell) than for classical improvement. For Wilson-type quark actions on an isotropic lattice, however, it is easy to see that the clover and Wilson terms are the only ones allowed at $\mathcal{O}(a)$ by gauge and (discrete) rotational symmetry [6]. Their coefficients of course renormalize on the quantum level. By a field transformation of the canonical form discussed above one can adjust the coefficient of one of these terms to any desired value. It is natural to do so for the Wilson term; in the SW case to maintain its “bare” coefficient

at the canonical value $r = 1$, for example. To eliminate quantum $\mathcal{O}(a)$ errors one then has to tune the coefficient of the clover term.

On an anisotropic lattice the situation is slightly more complicated. The allowed operators at $\mathcal{O}(a)$ consist of the spatial and temporal parts of the Wilson and clover terms, and the additional operator $[\gamma_0 D_0, \sum_i \gamma_i D_i]$. The most general field transformations Ω_c and $\bar{\Omega}_c$ allowed in this situation lead to three redundant operators, so that one has to tune two coefficients at $\mathcal{O}(a)$. These can be chosen to be the spatial and temporal parts of the clover term. Note that on an anisotropic lattice one must also allow a relative coefficient between the temporal and spatial kinetic terms at $\mathcal{O}(a^0)$, which can be tuned non-perturbatively by demanding a relativistic dispersion relation for the pion, say, at small masses and momenta.

3.3 The D234 Actions

Going to the next order in the expansion of the continuum derivatives in (3.5) gives the class of D234 actions

$$\begin{aligned} M_{\text{D234}} = & m_c(1 + \tfrac{1}{2}ra_0m_c) + \sum_{\mu} \gamma_{\mu} \nabla_{\mu} (1 - b_{\mu}a_{\mu}^2\Delta_{\mu}) \\ & - \tfrac{1}{2}ra_0 \left(\sum_{\mu} \Delta_{\mu} + \tfrac{1}{2}\sigma \cdot F \right) + \sum_{\mu} c_{\mu}a_{\mu}^3\Delta_{\mu}^2 \end{aligned} \quad (3.8)$$

where, at this point,

$$b_{\mu} = \frac{1}{6}, \quad c_{\mu} = \frac{ra_0}{24a_{\mu}} \quad (3.9)$$

The specific D234 action defined by the coefficients in (3.9) will be referred to as “D234c(r)”, where “c” refers to the fact that this action is obtained by our “canonical” field redefinition without any further modifications. If we use an improved representation of the field strength, as in eq. (B.16), this action only has $\mathcal{O}(a^4)$ classical errors. There is no canonical choice of r for this action. It will generically have three ghost branches, as illustrated in fig. 3 for the case of $r = 1$ (for $r = 2$ there are only two ghosts, but the lowest one is too low for this choice to be interesting, except perhaps on very anisotropic lattices).

It remains to be investigated what effect these ghost branches have on the quantum level, but one should certainly consider designing actions with fewer and higher-lying ghost branches. As we will see, the necessary “tuning” of the D234 actions introduces classical errors in addition to the $\mathcal{O}(a^4)$ ones. However, this is probably irrelevant, since the latter errors are unlikely to dominate on the quantum level anyhow.

3.4 Tuning the D234 Actions

To investigate the ghosts let us study the free dispersion relation corresponding to the general D234 action (3.8). The details are discussed in appendix C, to which we refer for the proof of any non-obvious facts we will use. For generic r, b_{μ} and c_{μ} the dispersion

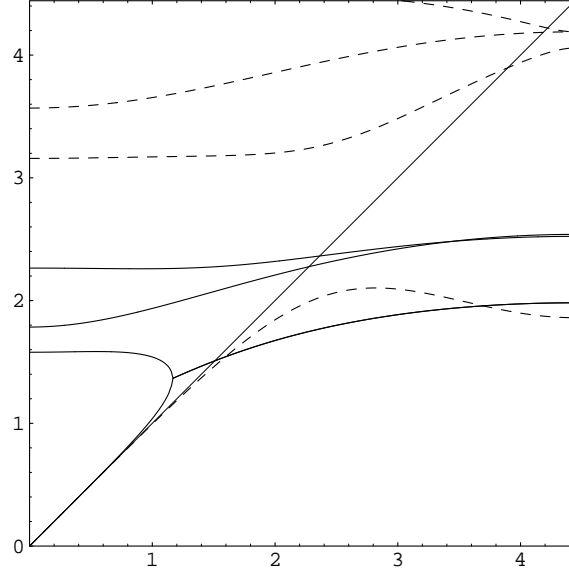


Figure 3: As in figure 2, for the massless D234c(1) action on 1:1 (solid) and 2:1 (dashed) lattices. We only show the real part of the energy, relevant for the top branch of each anisotropy, which has imaginary part π/a_0 , and to the right of the branch point on the 1:1 lattice.

relation is a quartic equation for $\sinh(a_0 E/2)^2$, $E = E(\mathbf{p})$, so there will be three ghost branches.⁵ Since the coefficients of the quartic are real, the energies will be real, come in complex conjugate pairs, or have imaginary part π/a_0 . Note that the qualitative branch structure (*e.g.* the number of branches) depends only on the temporal coefficients r, b_0 and c_0 . For $m_c = 0$ and $\mathbf{p} = \mathbf{0}$ the only way a lattice spacing enters the dispersion relation is via $a_0 E$. This implies that for small momenta and masses the height of the ghosts is inversely proportional to the temporal lattice spacing.

The most basic question we can ask about the ghosts, is how many ghost branches we can completely eliminate by a suitable choice of the free parameters. We summarize the conclusions of appendix C concerning this question as follows:

- If we choose $b_0 = 2c_0$ there will be at most two ghosts.
- If we further choose $r = 1 - 2b_0$ or $b_0 = 0$ there is (at most) one ghost.
- The only way to eliminate all ghosts is to choose $r = 1, b_0 = c_0 = 0$, which is of course the standard SW/Wilson case (if $r \neq 1$ the SW/Wilson action will have one ghost).

⁵Remember that we do not count the particle anti-particle symmetry $E \leftrightarrow -E$ in the number of solutions.

If we want to go beyond the SW action we will therefore have at least one ghost branch. Let us now discuss our “favorite” D234 actions on isotropic and anisotropic lattices in turn.

3.4.1 Isotropic D234 Action

On an isotropic lattice one will presumably prefer an isotropic action with manifest space-time exchange symmetry, so that one does not have to restore this symmetry by tuning on the quantum level. With $b_\mu = \frac{1}{6}$, the above results imply that when we require one ghost we must choose $r = \frac{2}{3}$ and $c_\mu = \frac{1}{12}$. We will refer to this action as the “isotropic D234” action. It was introduced in [10].

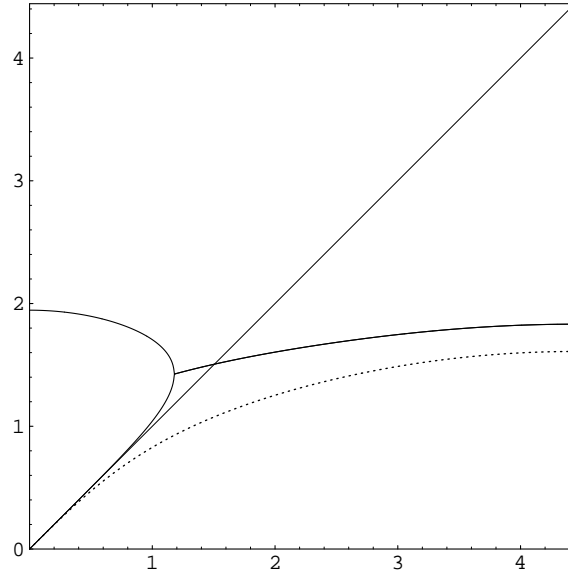


Figure 4: As in figure 3, for the massless isotropic D234 (solid) and SW (dotted) actions.

Since the c_μ violate eq. (3.9) this action has classical $\mathcal{O}(a^3)$ errors. At zero mass and momentum the one ghost branch of this action is at $aE(0) = \ln 7 \approx 1.9459$. Its massless dispersion relation is shown in figure 4, together with that of the SW/Wilson action.

We will see in sect. 3.5.3 that, like the SW and Wilson actions, this action can be coded very efficiently using the “projection trick”.

3.4.2 Anisotropic D234 Action

On an anisotropic lattice we can have one ghost with only $\mathcal{O}(a_0^3, a_\mu^4)$ errors simply by modifying the coefficient c_0 of the D234c($\frac{2}{3}$) action to be $c_0 = \frac{1}{12}$. We will refer to this

action by the name $D234i(\frac{2}{3})$.⁶ Its dispersion relation on a 2:1 lattice is compared in figure 5 with that of the SW action. Note in particular the impressive dispersion relation in the massive case, indicating that such an action might be very useful for heavy quark simulations.

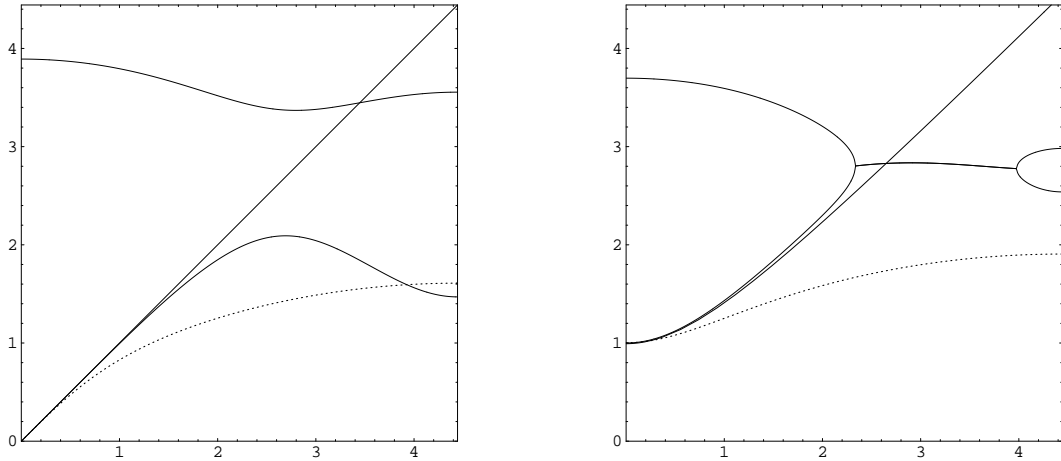


Figure 5: As in fig. 2, for $D234i(\frac{2}{3})$ on a 2:1 (solid) and SW/Wilson fermions on a 1:1 (dotted) lattice. The real part of $E(\mathbf{p})$ is shown to the right of the $D234i(\frac{2}{3})$ branch point. In the massive case the bare masses are tuned so that $a_s E(0) = 1$.

We should point out that the restriction of the $D234i(\frac{2}{3})$ action to an isotropic lattice does *not* give the isotropic D234 action. Because the spatial c_i of the former were chosen to not introduce a_i^3 errors, this action has anisotropic coefficients even on an isotropic lattice; it was designed for use on anisotropic lattices.

3.4.3 Variations

By relaxing the requirement of just one ghost one can construct actions that might be interesting for sufficiently anisotropic lattices. We will not discuss these here, but just make the general remark that for larger anisotropies a_s/a_t it is advantageous to choose larger values of r . Otherwise one will recover spatial doublers, as is obvious from the fact that our canonical field transformation (3.4) $\Omega_c \rightarrow 1$ as $a_0 \rightarrow 0$ for fixed r .⁷

We should briefly discuss one modification of the $D234i(\frac{2}{3})$ action. Namely, it is possible to construct a very similar D234 action that in the free case has only $\mathcal{O}(a^4)$ classical errors, with $\mathcal{O}(a_0^3)$ errors entering only in the presence of a gauge field.

⁶The “i” stands for mass-“independent”, since the coefficients in this action enjoy this property. In the next subsection we will describe a closely related D234 action with mass-dependent coefficients, to which we have previously [11] given the name $D234(\frac{2}{3})$.

⁷If one avoids the reappearance of doublers by letting ra_0 approach some non-zero limit as $a_0 \rightarrow 0$, one will obtain one ghost branch, with energy $2/ra_0$ for small masses and momenta.

This action is obtained by a somewhat more complicated change of variable from the continuum Dirac action,

$$\bar{\Omega}_c = \Omega_c, \quad \bar{\Omega}_c \Omega_c = 1 - \frac{1}{2} r a_0 (\not{D} - m_c - K_\epsilon), \quad K_\epsilon = \sum_\mu \epsilon_\mu a_\mu^2 \gamma_\mu D_\mu^3 \quad (3.10)$$

with free parameters ϵ_μ in addition to r . Requiring one ghost and $\mathcal{O}(a_0^3)$ errors only in the presence of a gauge field leads uniquely to the choice of coefficients

$$\begin{aligned} \epsilon_0 &= \frac{1}{6} \frac{1}{1 + \frac{3}{4} a_0 m_c} \\ b_0 &= \frac{1}{6} \frac{1 + \frac{1}{4} a_0 m_c}{1 + \frac{7}{12} a_0 m_c} \\ c_0 &= \frac{1}{2} b_0 \\ r &= \frac{2}{3} \frac{1 + \frac{3}{4} a_0 m_c}{1 + \frac{7}{12} a_0 m_c}, \end{aligned} \quad (3.11)$$

as well as $\epsilon_i = 0$, $b_i = \frac{1}{6}$ and $c_i = r a_0 / 24 a_i$. We will refer to this action as D234($\frac{2}{3}$) (labelling it by the value of r at $m_c = 0$). This was the action used in the anisotropic lattice simulations described in [11]. Note that numerically this action differs significantly from D234i($\frac{2}{3}$) only for very large masses; for $m_c = 0$ they are identical.

Finally, we remark that it is easy to improve the dispersion relation of D234-like actions at large momenta still further by introducing suitable fifth and sixth order derivative terms. This can be done by a field transformation and/or by simply adding such terms to the action. However, in the relevant momentum regime the hadron dispersion relations measured in simulations of D234 actions are already so good, even on coarse lattices [10, 11], that the additional cost and complication from the higher derivative terms seems unjustified.

3.5 Other Actions

We conclude this section by briefly discussing several other classes of actions (or other ways of writing actions) that are of interest for various conceptual and practical reasons.

3.5.1 Ghost-free D234-like Actions

We will now demonstrate that it is straightforward to write down a highly improved action, at tree level, that has no ghosts whatsoever. This comes at a price, of course. Such an action is more complicated and therefore more costly to simulate.

The idea is to use field transformations to eliminate the cubic temporal terms $\gamma_0 \nabla_0 \Delta_0$ in the naive improved fermion action in favor of spatial terms. Starting with the continuum action $M_c = m_c + \not{D}$, we perform a field transformation $\psi_c = \Omega_{c1} \psi$, $\bar{\psi}_c = \bar{\psi} \bar{\Omega}_{c1}$

with⁸

$$\begin{aligned}\Omega_{c1} &= 1 + \frac{a_0^2}{12} \left[D_0^2 - (\not{D} - m_c)(\not{D} + m_c) \right] \\ \bar{\Omega}_{c1} &= 1 + \frac{a_0^2}{12} \left[D_0^2 - (\not{D} + m_c)(\not{D} - m_c) \right],\end{aligned}\quad (3.12)$$

where the purely spatial derivative $\not{D} = \sum_i \gamma_i D_i$. This gives

$$\begin{aligned}\bar{\Omega}_{c1} M_c \Omega_{c1} &= m_c + \not{D} + \frac{1}{6} a_0^2 \gamma_0 D_0^3 + \delta K_c + \mathcal{O}(a^4), \\ \delta K_c &= \frac{a_0^2}{12} \{ m_c^2 - \sum_i D_i^2 - \frac{1}{2} \sigma \cdot F, \not{D} + m_c \}\end{aligned}\quad (3.13)$$

Note that, when discretized, the term $\not{D} + \frac{1}{6} a_0^2 \gamma_0 D_0^3$ will not contain any lattice time derivative above the first, and there is no other temporal derivative in the action. We can now proceed with a second change of variable to introduce even derivatives, defined by $\bar{\Omega}_{c2} = \Omega_{c2}$ and

$$\bar{\Omega}_{c2} \Omega_{c2} = 1 + \frac{1}{2} r a_0 \left(\not{D} + \frac{1}{6} a_0^2 \gamma_0 D_0^3 - m_c - \delta K_c \right). \quad (3.14)$$

This implies

$$\begin{aligned}\bar{\Omega}_{c2} \bar{\Omega}_{c1} M_c \Omega_{c1} \Omega_{c2} &= m_c (1 + \frac{1}{2} r a_0 m_c) + \not{D} + \frac{1}{6} a_0^2 \gamma_0 D_0^3 + (1 + r a_0 m_c) \delta K_c \\ &\quad - \frac{1}{2} r a_0 (\not{D} + \frac{1}{6} a_0^2 \gamma_0 D_0^3)^2 + \mathcal{O}(a^4).\end{aligned}\quad (3.15)$$

Finally, we discretize this action. Using

$$(\not{D} + \frac{1}{6} a_0^2 \gamma_0 D_0^3)^2 = \sum_\mu \Delta_\mu - \frac{1}{12} \sum_i a_i^2 \Delta_i^2 + \frac{1}{2} \sigma \cdot F + \mathcal{O}(a_0^2, a^4), \quad (3.16)$$

we find

$$\begin{aligned}M_{\text{D234gf}} &= m_c (1 + \frac{1}{2} r a_0 m_c) + \nabla_0 + \nabla_c + (1 + r a_0 m_c) \delta K \\ &\quad - \frac{1}{2} r a_0 \left(\sum_\mu \Delta_\mu + \frac{1}{2} \sigma \cdot F \right) + \frac{r a_0}{24} \sum_i a_i^2 \Delta_i^2,\end{aligned}\quad (3.17)$$

$$\delta K = \frac{a_0^2}{12} \{ m_c^2 - \sum_i \Delta_i - \frac{1}{2} \sigma \cdot F, \nabla + m_c \} = \delta K_c + \mathcal{O}(a^4), \quad (3.18)$$

where $\nabla_0 = \gamma_0 \nabla_0$ is the unimproved temporal lattice derivative, ∇_c is the improved spatial derivative (see (B.12)) and for the field strength F one should use an improved discretization. This action has $\mathcal{O}(a_0^3, a^4)$ errors in on-shell quantities. As usual, by

⁸This field transformation is similar to ones used in [21].

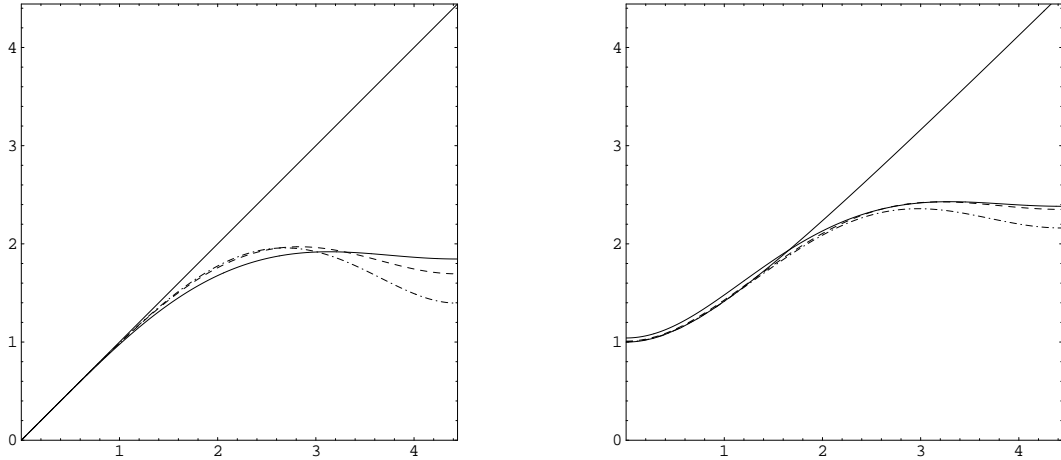


Figure 6: As in figure 2, for the ghost-free D234-like action on 1:1 (solid), 2:1 (dashed) and 3:1 (dot-dashed) lattices.

undoing the above field transformations one can achieve the same errors for off-shell quantities.

Obviously the above action is the analog of the D234 actions of the previous sections, with δK appearing in place of cubic and quartic temporal derivative terms. Now, however, we have just one branch if we set $r = 1$, as for the Wilson and SW actions. This is illustrated in figure 6.

For the a_0^3 errors to be negligible compared to the a^4 ones, we can again use anisotropic lattices. In that case the ghost branches of the D234 actions of the previous section are presumably harmless, and it seems doubtful that having no ghosts outweighs the disadvantage of having to include the costly anti-commutator term δK . Comparisons to simulations with this action should, however, allow one to discern whether the ghost branches have any effect besides that on correlation functions at small times.

Another, less ambitious ghost-free action can easily be constructed if one is willing to tolerate a_0^2 in addition to a_0^3 and a^4 errors. Such an action can be obtained, for example, by simply neglecting the temporal third and fourth order terms in the D234c(1) action.

3.5.2 D34 Action

Recently a tadpole-improved version of an improved action discussed in [29] was used in a Monte Carlo simulation of the hadron spectrum [30]. This action has third and fourth order derivative terms, but no second order Wilson term. We will therefore refer to it as the D34 action. Generalized to an anisotropic lattice it reads in our notation

$$M_{\text{D34}} = m_c + \nabla_c + \sum_{\mu} c_{\mu} a_{\mu}^3 \Delta_{\mu}^2. \quad (3.19)$$

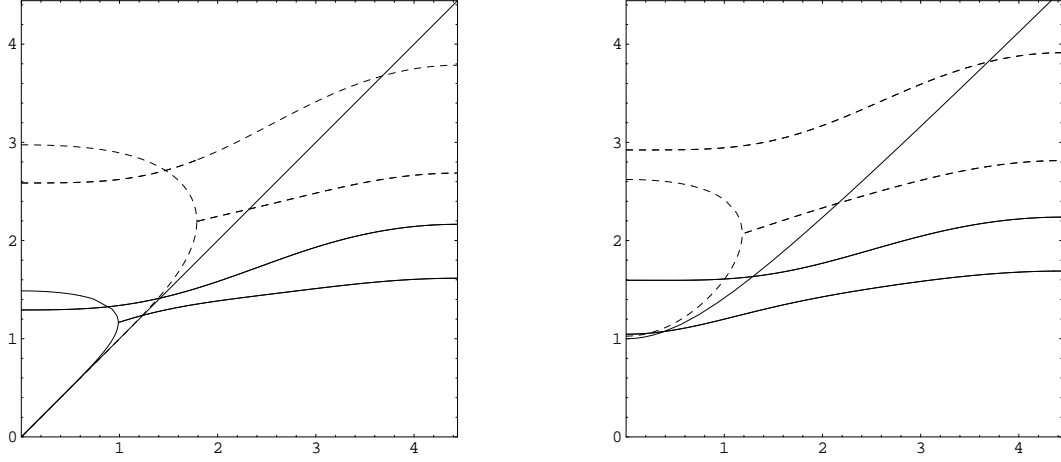


Figure 7: As in figure 5, for D34 fermions with $c_\mu = \frac{1}{6}$ on 1:1 (solid) and 2:1 (dashed) lattices. Note that this action has four branches. They consist of two pairs of complex conjugate solutions for sufficiently large momenta. For small momenta there are usually two real branches. For $a_s m_c = 1$ on an isotropic lattice, however, all branches are complex even at small momenta.

In [30] an isotropic lattice with all $c_\mu = \frac{1}{6}$ was used. The D34 action is obtained from the naive improved fermion action not by a change of variable, but simply by adding the Δ^2 term, which leads to an a^3 error (and means that the fields are already improved up to a^3 errors; no change of variable has to be undone). This action is of course a special case of the general class of D234 actions, with $r = 0$ and $b_\mu = \frac{1}{6}$. It has three ghosts, except for $c_0 = \frac{1}{12}$, when there are two. There is no obvious canonical choice for the c_μ ; for all values of c_μ the ghost branches seem to lie rather low. In figure 7 we show various dispersion relations for the $c_\mu = \frac{1}{6}$ case.

3.5.3 W-Actions

Recall that the Wilson and SW actions with $r = 1$ can be efficiently coded using the “projection trick”, which exploits the fact that the “Wilson operator” [31]

$$W_\mu \equiv \nabla_\mu - \frac{a_\mu}{2} \gamma_\mu \Delta_\mu \quad (3.20)$$

can be expressed in terms of (spinor-) projection operators. Most conveniently this is expressed as

$$\gamma_\mu W_\mu = -\nabla_\mu^+ P_\mu^- + \nabla_\mu^- P_\mu^+, \quad (3.21)$$

where $P_\mu^\pm \equiv \frac{1}{2}(1 \pm \gamma_\mu)$ are projectors on two-dimensional, orthogonal subspaces (for fixed μ), and

$$\nabla_\mu^\pm \psi(x) \equiv \pm \frac{1}{a_\mu} \left(U_{\pm\mu}(x) \psi(x \pm \mu) - \psi(x) \right) \quad (3.22)$$

are forward and backward derivatives.

Actions that can be expressed as (low-order) polynomials in the W_μ ,

$$M = \sum_{\mu} \frac{1}{a_{\mu}} Q_{\mu}(a_{\mu} \gamma_{\mu} W_{\mu}) \quad (3.23)$$

(with the possible addition of mass and clover terms, cf. below) will be cheap to code. Another advantage of such an action is that one can immediately read off the number of branches of the dispersion relation: If $Q_0(x) = x + \dots$ is an n -th order polynomial, there are exactly n branches, counting possible degeneracies (the proof of this statement is left as an exercise to the reader). Note that if all $Q_{\mu}(x)$ are n -th order polynomials, the class of such “W23... n ” actions is a proper subset of the class of “D23... $2n$ ” actions considered previously, which generically have $2n$ branches.

It is interesting to ask when such W-actions are improved. We first remark that the continuum derivative can be written as

$$a_{\mu} \gamma_{\mu} D_{\mu} = -\ln(1 - a_{\mu} \gamma_{\mu} W_{\mu}) = a_{\mu} \gamma_{\mu} W_{\mu} + \frac{1}{2}(a_{\mu} \gamma_{\mu} W_{\mu})^2 + \frac{1}{3}(a_{\mu} \gamma_{\mu} W_{\mu})^3 + \dots \quad (3.24)$$

Truncation of this expansion does, however, not seem to lead to promising (on- and off-shell) improved actions, since the expansion converges too slowly and the ghost branches lie too low.

Alternatively, we may ask when a W-action is only on-shell improved, related to some order in a by our canonical field transformation (3.4) to the continuum Dirac action. Restricting ourselves to isotropic actions from now on — the improved actions in question are not simply expressed in terms of W_{μ} on anisotropic lattices — we write

$$M = m_c(1 + \frac{1}{2}ram_c) + \frac{1}{a} \sum_{\mu} Q(a\gamma_{\mu}W_{\mu}) - \frac{1}{4}ra\sigma \cdot F. \quad (3.25)$$

For a first-order polynomial, $Q(x) = x$, we recover, of course, the SW action for $r=1$, with $\mathcal{O}(a^2)$ errors. To see which higher order polynomials correspond to improved actions, we expand out the W_{μ} in terms of ∇_{μ} and Δ_{μ} and compare with the results of the previous subsections (specifically, the D234c(r) actions and their higher order analogs).

The second order polynomial $Q(x) = x + \frac{1}{6}x^2$ with $r = \frac{2}{3}$ gives a “W2” action, which has $\mathcal{O}(a^3)$ errors. It is identical to the isotropic D234 action [10] of sect. 3.4.1. This implies that the application of the isotropic D234 operator is only about twice as expensive as that of the $r=1$ SW or Wilson operator.

For a third order polynomial one can reduce the errors to $\mathcal{O}(a^4)$ by choosing $Q(x) = x + \frac{5}{22}x^2 + \frac{2}{33}x^3$ and $r = \frac{6}{11}$ to give a “W23” action (one must now also use an improved F in the clover term). This action is equal to a D23456 action with $r = \frac{6}{11}$, $b_{\mu} = \frac{1}{6}$, $c_{\mu} = \frac{1}{44}$, $d_{\mu} = \frac{2}{33}$, $e_{\mu} = \frac{1}{33}$, where d_{μ} and e_{μ} are the coefficients of fifth, respectively, sixth order derivative terms, defined analogously to b_{μ} and c_{μ} . This action has two ghost branches; for all masses and momenta they form a complex-conjugate pair. For zero mass and momentum their energies are $aE(0) \approx 1.528 \pm 0.897i$.

4 Large Mass Behavior of the D234 Actions

Our D234 actions were initially designed with light quarks in mind. In contrast to [21], for example, where renormalization conditions are imposed on mass-shell, even for heavy quarks, we effectively do so at zero quark mass, since we always expand in powers of a_μ , assuming $a_\mu p_\mu$ to be small. Heavy onium systems are non-relativistic, so masses are large but momenta are small. One would therefore expect that by going to an anisotropic lattice, where $a_t m_c$ is small, our D234 actions could be used to simulate such systems on lattices where the spatial lattice spacing is still quite large. Decreasing the temporal lattice spacing by a factor of, say, 3 – 5, would present a relatively minor increase in cost, given the exciting prospect of obtaining accurate results for *e.g.* charmonium within a relativistic framework. The charmonium system is difficult to simulate. It is light enough for the NRQCD expansion to become problematic, but too heavy for the usual light quark actions on isotropic lattices to be accurate. A highly improved quark action on an anisotropic lattice seems tailor-made to simulate such a system.

To get a quantitative idea of how small we have to choose a_t , let us investigate in the free case when the D234 actions break down at large masses. As indicators of break-down we will consider $E(0)/m_c$ and the “effective velocity of light” $c(\mathbf{p})$, defined by

$$c(\mathbf{p})^2 = \frac{E(\mathbf{p})^2 - E(0)^2}{\mathbf{p}^2}, \quad (4.1)$$

where $E(\mathbf{p})$ is meant to be the physical branch of the energy. We know that for small masses $E(0)/m_c$ and $c(0)$ are 1 up to order $\mathcal{O}(a_t^3 m_c^3)$ or $\mathcal{O}(a_t^4 m_c^4)$ corrections. For large masses, it is clear from the figures presented in sect. 3, that for the relevant D234 actions the two lowest real branches of the energy (at fixed momentum) will eventually merge. (For zero momentum this branch point occurs at the same value of $a_t m_c$ for any anisotropy. For non-zero momentum this becomes true only asymptotically, for large anisotropy.) At the branch point $c(\mathbf{p})$ diverges.

In figure 8 we compare $c(\mathbf{p})$ for the D234i($\frac{2}{3}$) and SW/Wilson fermions for various anisotropies. We show $c(\mathbf{p})$ as a function of mass for both $\mathbf{p} = \mathbf{0}$ and the (quite large) value $\mathbf{p} = (1, 0, 0)/a_s$.

As expected, $c(\mathbf{p})$ stays close to 1 for a larger and larger mass range as the anisotropy increases. Note how much better the D234 action behaves in this respect than the SW/Wilson action, in particular at non-zero momentum. It is also interesting to observe that for the SW/Wilson action $c(\mathbf{p})$ decreases more or less monotonically as a function of both m_c and $|\mathbf{p}|$, whereas for the D234 action $c(\mathbf{p})$ increases as a function of m_c for fixed \mathbf{p} , but (more or less) decreases as a function of $|\mathbf{p}|$ for fixed m_c (an isotropic lattice provides an exception to this latter statement). This is presumably one of the reasons we find in Monte Carlo simulations that the D234 actions exhibit an excellent dispersion relation for various hadrons up to surprisingly large masses and momenta.

As is clear from figure 8, for the D234 actions the transition from $c(\mathbf{p}) \approx 1$ to divergence is quite rapid. This provides a sensitive indicator of the breakdown of the D234 actions at large masses. When $c(\mathbf{p})$ begins to get larger than 1, the D234 actions

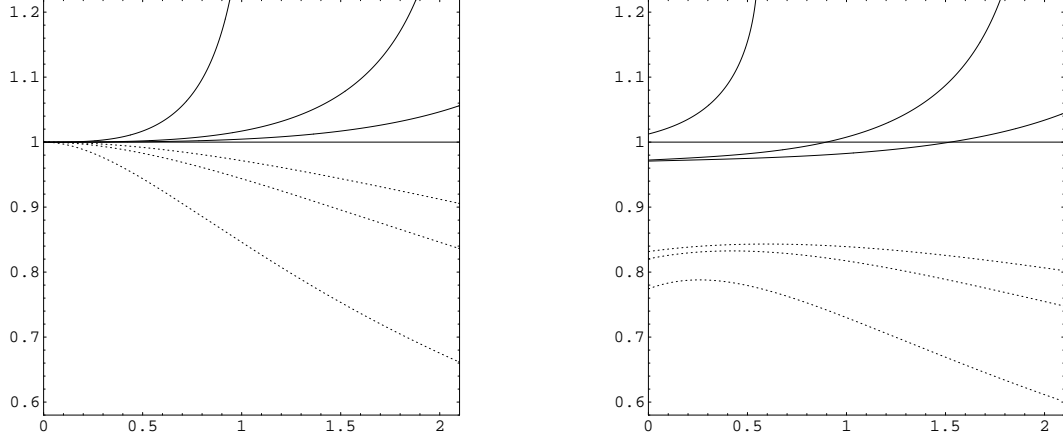


Figure 8: $c(\mathbf{p}=\mathbf{0})$ (left) and $c(\mathbf{p}=(1,0,0)/a_s)$ (right) as functions of $a_s m_c$ for D234i($\frac{2}{3}$) (solid) and SW/Wilson (dotted) fermions on 1:1, 2:1 and 3:1 lattices.

should not be trusted anymore. When this happens, it should only be necessary to decrease a_t by a relatively modest amount to be able to simulate the quark mass of interest.

To reinforce this last point, let us mention that among the anisotropic D234 actions we discussed, the finite-mass errors are largest for the D234i($\frac{2}{3}$) action, where

$$\frac{E(0)}{m_c} = 1 + \frac{1}{18} a_t^3 m_c^3 + \mathcal{O}(a_t^4 m_c^4) \quad (4.2)$$

$$c(0) = 1 + \frac{1}{9} a_t^3 m_c^3 + \mathcal{O}(a_t^4 m_c^4) . \quad (4.3)$$

For comparison, the corresponding expansions for the free SW/Wilson action are

$$\begin{aligned} \frac{E(0)}{m_c} &= 1 - \frac{1}{6} a_t^2 m_c^2 + \mathcal{O}(a_t^3 m_c^3) \\ c(0) &= 1 - \frac{1}{3} a_t^2 m_c^2 + \mathcal{O}(a_t^3 m_c^3) . \end{aligned} \quad (4.4)$$

We see that not only do the D234 actions have a larger power in their scaling errors, they are also blessed with small coefficients. The coefficients in the above expansions for various other actions can be found in appendix C. There we also discuss the error that arises in the presence of a non-zero gauge field. Again we find that the coefficients of the error terms are very small as long as $a_t m_c \leq \mathcal{O}(1)$.

For the large mass error in $c(0)$ to stay below the one or two percent level, one must choose $a_t m_c < 0.5$ for the D234 actions. For the SW action one must satisfy the much tighter constraint $a_t m_c < 0.2$ to achieve the same error. Given this bound for the D234

actions, we can attain a meson mass

$$m_{q\bar{q}} \approx 2m_c = \frac{2\xi}{a_s} a_t m_c \quad (4.5)$$

in the charmonium range ($m_{q\bar{q}} \approx 3.0 \text{ GeV}$) with a spatial lattice spacing of $a_s^{-1} \approx 600 - 1000 \text{ MeV}$ and anisotropy $\xi = 3 - 5$. For the SW action the lattice spacings would have to be almost three times smaller with the same anisotropy, making a Monte Carlo simulation roughly two orders of magnitude more expensive. Needless to say, on an isotropic lattice a simulation of charmonium in a relativistic formalism would be more expensive by many, many additional orders of magnitude.

5 Discussion and Conclusions

We have seen that using field transformations it is easy to design highly improved, doubler-free quark actions at tree level. We emphasize that the improvement includes all interactions between quarks and gluons, not just the free quark dispersion relation. We have discussed several actions, representing different compromises between the conflicting aims of high level of improvement, absence of ghost branches, and simplicity. The actions we are currently using in simulations are the isotropic D234 (sect. 3.4.1) and the anisotropic D234i($\frac{2}{3}$) (sect 3.4.2) actions. However, given the ingredients provided, the reader can concoct many more flavors of actions.

The actions we constructed are just moderately more expensive to simulate than the SW action. For example, the isotropic D234 action is only about twice, the generic D234 action about four times as expensive than the isotropic SW action with $r = 1$ (a slight additional overhead might be incurred in preconditioning).

The next task, of course, is to put some quantum flesh on these classical bones. Exploratory quenched simulation results using tadpole improvement have appeared [10, 11], and are very encouraging. Further work is in progress [22]. In terms of the general improvement program, the next step is to check if a non-perturbative tuning of the low-order coefficients in the action is necessary. To eliminate all errors up to $\mathcal{O}(a)$, one would only have to tune the clover term, for an isotropic action. It was recently shown [7] how to implement this in practice, by demanding the restoration of chiral symmetry at zero quark mass. In [7] the case of SW quarks on Wilson glue was considered, but the methods apply equally well to other actions.

As mentioned in the introduction (cf. sect. 3.2 for more details), for anisotropic actions there are more coefficients that can renormalize on the quantum level, and in principle have to be tuned in order to restore space-time exchange symmetry. We hope that suitable tadpole improvement on anisotropic lattices [18] will reduce these renormalizations so that their effect on physical quantities is only on the few percent level, where they can be neglected; but this is an issue to be decided by empirical study.

Already at $\mathcal{O}(a^0)$ one might have to introduce a “bare velocity of light” into the action, to restore space-time exchange symmetry at leading order. The required renormalization is easily determined non-perturbatively in this case, by measuring

the dispersion relation of a pion, say, at small masses and momenta. For on-shell improvement at $\mathcal{O}(a)$ there is one additional coefficient compared to the isotropic case, which one can choose to be the relative coefficient of the temporal and spatial parts of the clover term. Using the methods of [7] with a suitable background field, for example, it should be possible to determine this renormalization. Finally, we should point out that on an anisotropic lattice the effect of these renormalizations is suppressed, since the overall coefficient of the $\mathcal{O}(a)$ terms is smaller by a factor a_t/a compared to the isotropic case (for fixed r).

In summary, the D234 actions show great promise for accurate QCD simulations on coarse lattices of improved glue, for both light and heavy quarks. The low-order terms in the actions require further study, and, if necessary, should be determined non-perturbatively. We hope that, as for improved glue, quark actions with negligible $\mathcal{O}(a)$ and only small quantum $\mathcal{O}(a^2)$ errors will give accurate results for *all* observables on coarse lattices. The next task of our program will be the determination of renormalization constants like Z_A and Z_V for the D234 actions. Once quenched QCD is satisfactorily understood, it is time for simulations of full QCD. Realistic simulations of full QCD should finally be possible on coarse lattices.

Acknowledgements

We would like to thank Robert Edwards, Urs Heller, Tony Kennedy, Paul Mackenzie, Stefan Sint, and John Sloan for discussions. This work is supported by DOE grants DE-FG05-85ER25000, DE-FG05-92ER40742 and DE-FG02-90ER40542.

Appendix A Euclidean Continuum QCD

We write the action of euclidean $SU(N)$ gauge theory in four dimensions as

$$S_g[A] = \frac{1}{2g^2} \int d^4x \operatorname{Tr} F_{\mu\nu}(x) F_{\mu\nu}(x) , \quad (\text{A.1})$$

where $F_{\mu\nu}(x)$ is the $\mathfrak{su}(N)$ -valued hermitean field strength. In our conventions the covariant derivative is written in terms of the hermitean gauge field $A_\mu(x)$ as

$$D_\mu = \partial_\mu - iA_\mu \quad (\text{A.2})$$

so that

$$F_{\mu\nu} = i[D_\mu, D_\nu] = \partial_\mu A_\nu - \partial_\nu A_\mu - i[A_\mu, A_\nu] . \quad (\text{A.3})$$

In terms of traceless hermitean $\mathfrak{su}(N)$ generators T_a , $a = 1, \dots, N^2 - 1$, normalized by $\operatorname{Tr}(T_a T_b) = \frac{1}{2} \delta_{ab}$, we write

$$F_{\mu\nu}(x) = F_{\mu\nu}^a(x) T_a, \quad A_\mu(x) = A_\mu^a(x) T_a . \quad (\text{A.4})$$

The parallel transporter from a point x to y along a curve \mathcal{C}_{yx} is

$$U[\mathcal{C}_{yx}] = \text{P exp} \left(i \int_{\mathcal{C}_{yx}} dx'_\mu A_\mu(x') \right) \in SU(N) , \quad (\text{A.5})$$

where P denotes path ordering, stipulating that in a series or product expansion of the exponential the fields at a point earlier on the curve are to be placed to the right of fields at later points. Under a local gauge transformation, $\Lambda(\cdot) \in SU(N)$,

$$\begin{aligned} F_{\mu\nu}(x) &\rightarrow \Lambda(x) F_{\mu\nu}(x) \Lambda^{-1}(x) \\ U[\mathcal{C}_{yx}] &\rightarrow \Lambda(y) U[\mathcal{C}_{yx}] \Lambda^{-1}(x) . \end{aligned} \quad (\text{A.6})$$

Recall that the inhomogenous transformation law of $A_\mu(x)$ has been *designed* so that the parallel transporter transforms covariantly as above.

The action of a Dirac fermion coupled to a gauge field is

$$S_f[\psi, \bar{\psi}] = \int d^4x \bar{\psi}(x) (\not{D} + m) \psi(x) , \quad (\text{A.7})$$

where $\not{D} = \sum_\mu \gamma_\mu D_\mu$ in terms of the euclidean gamma matrices defined by

$$\begin{aligned} \gamma_\mu &= \gamma_\mu^\dagger \\ \{\gamma_\mu, \gamma_\nu\} &= 2\delta_{\mu\nu} . \end{aligned} \quad (\text{A.8})$$

The spinor fields $\psi(x)$ and $\bar{\psi}$ also carry a suppressed color index in the vector representation of $SU(N)$. Under a gauge transformation

$$\begin{aligned} \psi(x) &\rightarrow \Lambda(x) \psi(x) \\ \bar{\psi}(x) &\rightarrow \bar{\psi}(x) \Lambda^{-1}(x) \end{aligned} \quad (\text{A.9})$$

and $D_\mu\psi(x)$ transforms like $\psi(x)$ (again, by construction of $A_\mu(x)$).

We will need some identities involving the hermitean $\sigma_{\mu\nu}$ matrices defined by

$$\sigma_{\mu\nu} = -\frac{i}{2}[\gamma_\mu, \gamma_\nu] = -\sigma_{\nu\mu}, \quad \text{or} \quad \gamma_\mu\gamma_\nu = \delta_{\mu\nu} + i\sigma_{\mu\nu}. \quad (\text{A.10})$$

This implies

$$\frac{1}{2}\{\gamma_\mu D_\mu, \gamma_\nu D_\nu\} = \delta_{\mu\nu} D_\mu^2 + \frac{1}{2}\sigma_{\mu\nu} F_{\mu\nu} \quad (\text{A.11})$$

Appendix B Lattice QCD

We will consider a four-dimensional euclidean hypercubic lattice of extent L_μ and lattice spacing a_μ in direction $\mu = 0, 1, 2, 3$. Since we will always work in euclidean space (except when we consider dispersion relations) the use of $\mu=0$ to denote the euclidean time direction should not cause confusion.

Points are labelled by x, y, \dots , as in the continuum. When all spatial lattice spacings are identical, they are denoted by a_s ; the temporal lattice spacing will then be denoted by $a_t = a_0$. We will refer to $\xi \equiv a_s/a_t$ as the *anisotropy* of the lattice, and sometimes call such a lattice a “ $\xi : 1$ lattice”. For an isotropic lattice we set $a \equiv a_\mu$.

We use i, j, \dots for spatial indices, and boldface letters for spatial vectors. The notation $x \pm \mu$ is a shorthand for $x \pm a_\mu \hat{\mu}$, where $\hat{\mu}$ is a unit vector in the positive μ -direction.

When working on anisotropic lattices one should in principle be very careful in specifying the lattice spacing errors of various quantities. To avoid cumbersome notation, we will be careful only to distinguish errors of the form $\mathcal{O}(a_\mu^n)$ from $\mathcal{O}(a^n)$ errors, where the latter denotes any errors that are not of the form $\mathcal{O}(a_\mu^n)$ or are a sum of such terms with different μ .

Gauge field dynamics on a lattice is expressed in terms of the link field $U_\mu(x)$, which takes values in the gauge group. $U_\mu(x)$ is a parallel transporter from $x + \mu$ to x , along a straight line. In terms of an underlying continuum gauge field $A_\mu(x)$ we therefore have (cf. appendix A)

$$U_\mu(x) = \text{P exp} \left(-i \int_x^{x+\mu} dx'_\mu A_\mu(x') \right), \quad (\text{B.1})$$

where now (because we changed the sign of the path compared to eq. (A.5)) fields at points *later* on the path are to be placed to the right of earlier fields. We will employ the notation $U_{-\mu}(x) \equiv U_\mu(x - \mu)^\dagger$ for the parallel transporter from $x - \mu$ to x .

Using the link field $U_\mu(x)$ is the only known way of maintaining exact gauge invariance on the lattice (not preserving manifest gauge invariance would necessitate a costly tuning of the various gauge couplings). Under a gauge transformation

$$U_{\pm\mu}(x) \rightarrow \Lambda(x) U_\mu(x) \Lambda^{-1}(x \pm \mu). \quad (\text{B.2})$$

With the help of the link field it is trivial to construct gauge-covariant first- and second-order lattice derivatives via

$$\nabla_\mu \psi(x) \equiv \frac{1}{2a_\mu} \left[U_\mu(x) \psi(x + \mu) - U_{-\mu}(x) \psi(x - \mu) \right] \quad (\text{B.3})$$

$$\Delta_\mu \psi(x) \equiv \frac{1}{a_\mu^2} \left[U_\mu(x) \psi(x + \mu) + U_{-\mu}(x) \psi(x - \mu) - 2\psi(x) \right]. \quad (\text{B.4})$$

The operators ∇_μ and Δ_μ are the building blocks of our lattice fermion actions.

It is useful to observe that in momentum space (with $U_\mu \equiv 1$) they correspond to

$$\nabla_\mu \leftrightarrow i\bar{p}_\mu, \quad a_\mu \bar{p}_\mu \equiv \sin(a_\mu p_\mu) \quad (\text{B.5})$$

$$\Delta_\mu \leftrightarrow -\hat{p}_\mu^2, \quad a_\mu \hat{p}_\mu \equiv 2 \sin(a_\mu p_\mu/2) \quad (\text{B.6})$$

Momenta on a lattice can be taken to lie in the Brillouin zone, defined by $-\pi/a_\mu < p_\mu \leq \pi/a_\mu$. Note that \bar{p}_μ , in contrast to \hat{p}_μ , has a “doubler problem”, *i.e.* it vanishes at the edge of the Brillouin zone $a_\mu p_\mu = \pi$.

Since ∇_μ and Δ_μ are gauge-covariant, many identities involving these operators are easy to prove by going to a gauge in which $U_\mu \equiv 1$, where such identities reduce to relations between \bar{p}_μ and \hat{p}_μ . For example,

$$\nabla_\mu \Delta_\mu = \Delta_\mu \nabla_\mu, \quad (\text{B.7})$$

and

$$\nabla_\mu \nabla_\mu = \Delta_\mu + \frac{1}{4} a_\mu^2 \Delta_\mu \Delta_\mu. \quad (\text{B.8})$$

Other identities we will need follow from

$$p_\mu = \bar{p}_\mu \left(1 + \frac{a_\mu^2}{6} \hat{p}_\mu^2 + \frac{a_\mu^4}{30} \hat{p}_\mu^4 + \frac{a_\mu^6}{140} \hat{p}_\mu^6 + \dots \right) \quad (\text{B.9})$$

$$p_\mu^2 = \hat{p}_\mu^2 + \frac{a_\mu^2}{12} \hat{p}_\mu^4 + \frac{a_\mu^4}{90} \hat{p}_\mu^6 + \frac{a_\mu^6}{560} \hat{p}_\mu^8 + \dots \quad (\text{B.10})$$

The operator ∇_μ has $\mathcal{O}(a_\mu^2)$ errors compared to the continuum derivative,

$$\nabla_\mu = D_\mu + \mathcal{O}(a_\mu^2). \quad (\text{B.11})$$

Eq. (B.9) implies that a more continuum-like first-order covariant derivative, involving only next-nearest neighbor sites, can be defined as follows

$$\nabla_{c\mu} \equiv \nabla_\mu \left(1 - \frac{1}{6} a_\mu^2 \Delta_\mu \right) = D_\mu + \mathcal{O}(a_\mu^4). \quad (\text{B.12})$$

The clover operator is a well-known lattice representation of the field strength that agrees with the continuum $F_{\mu\nu}$ up to $\mathcal{O}(a^2)$ errors. To define it, let us expand on the path notation used for the parallel transporter in appendix A. Namely, let us introduce a shorthand for lattice paths, defined recursively so that $\mathcal{C}_{yx}(\mu)$ denotes the path from

$x + \mu$ to x to y , and $x = \mathcal{C}_{xx}$. Path ordering implies that $U[\mathcal{C}_{yx}(\mu)] = U[\mathcal{C}_{yx}]U_\mu(x)$. Of course, $U[\mathcal{C}_{xx}] = 1$.

The clover operator we use, $F_{\mu\nu}^{(cl)}(x)$, can be expressed in terms of the sum of the link fields around the four plaquettes bordering on x (all counter-clockwise, say), as

$$F_{\mu\nu}^{(cl)}(x) \equiv \frac{1}{4a_\mu a_\nu} \mathcal{T} \left(U[x(\mu)(\nu)(-\mu)(-\nu)] + U[x(\nu)(-\mu)(-\nu)(\mu)] + \right. \\ \left. U[x(-\mu)(-\nu)(\mu)(\nu)] + U[x(-\nu)(\mu)(\nu)(-\mu)] \right), \quad (\text{B.13})$$

where $\mathcal{T}(M)$ is the (color-)traceless imaginary part of an $N \times N$ matrix,

$$\mathcal{T}(M) \equiv \frac{1}{2}(M - M^\dagger) - \frac{i}{N} \text{ImTr } M. \quad (\text{B.14})$$

A simple calculation shows that

$$F_{\mu\nu}^{(cl)}(x) = F_{\mu\nu}(x) + \frac{1}{6}(a_\mu^2 D_\mu^2 + a_\nu^2 D_\nu^2) F_{\mu\nu}(x) + \mathcal{O}(a^4). \quad (\text{B.15})$$

Using a suitable discretization of the second-order derivatives acting on the field strength, we can therefore define a more continuum-like lattice field strength via

$$F_{c\mu\nu}(x) \equiv \frac{5}{3} F_{\mu\nu}^{(cl)}(x) - \frac{1}{6} \left[U_\mu(x) F_{\mu\nu}^{(cl)}(x + \mu) U_{-\mu}(x + \mu) + \right. \\ \left. U_{-\mu}(x) F_{\mu\nu}^{(cl)}(x - \mu) U_\mu(x - \mu) - (\mu \leftrightarrow \nu) \right] \\ = F_{\mu\nu}(x) + \mathcal{O}(a^4). \quad (\text{B.16})$$

Appendix C Dispersion Relation of the D234 Actions

In terms of \bar{p} and \hat{p} introduced in appendix B let us define

$$\tilde{p}_\mu \equiv \bar{p}_\mu (1 + b_\mu a_\mu^2 \hat{p}_\mu^2) \\ \tilde{m}(p) \equiv m_0 + \frac{1}{2} r a_0 \sum_\mu \hat{p}_\mu^2 + \sum_\mu c_\mu a_\mu^3 \hat{p}_\mu^4 \\ m_0 \equiv m_c (1 + \frac{1}{2} r a_0 m_c). \quad (\text{C.1})$$

The inverse of the free propagator of the general D234 action (3.8) is then simply $i\vec{p} + \tilde{m}(p)$ in momentum space, implying the dispersion relation

$$\tilde{p}^2 + \tilde{m}(p)^2 = 0. \quad (\text{C.2})$$

As a dimensionless measure of the energy it is useful to introduce

$$y \equiv y(E) \equiv -a_0^2 \hat{p}_0^2 = 4 \sinh^2\left(\frac{a_0 E}{2}\right), \quad (\text{C.3})$$

so that

$$\begin{aligned} -a_0^2 \tilde{p}_0^2 &= y \left(1 + \frac{y}{4}\right) \\ -a_0^2 \tilde{p}_0^2 &= y \left(1 + \frac{y}{4}\right) (1 - b_0 y)^2 . \end{aligned} \quad (\text{C.4})$$

Expressed in terms of y the dispersion relation for the energy $E = E(\mathbf{p})$ therefore reads

$$y \left(1 + \frac{y}{4}\right) (1 - b_0 y)^2 = a_0^2 \tilde{\mathbf{p}}^2 + \left(\mu(\mathbf{p}) - \frac{r}{2} y + c_0 y^2\right)^2 , \quad (\text{C.5})$$

where

$$\mu(\mathbf{p}) \equiv a_0 m_0 + \frac{1}{2} r a_0^2 \sum_i \hat{p}_i^2 + a_0 \sum_i c_i a_i^3 \hat{p}_i^4 . \quad (\text{C.6})$$

Note that the use of y factors out the particle anti-particle symmetry $E \leftrightarrow -E$, since y is invariant under it. The dispersion relation is generically a quartic equation for y with real coefficients. This implies that generically there will be four solutions for $\pm E$, which, at a given \mathbf{p} , are either real, come in complex conjugate pairs, or have imaginary part π/a_0 .

Let us now discuss when the quartic reduces to some lower order polynomial. We see immediately that the quartic term cancels when $c_0 = \frac{1}{2}b_0$. Assuming this holds, the cubic term cancels if $b_0=0$ or $r=1-2b_0$. The remaining quadratic equation reads

$$-y^2 \left[b_0(2 + \mu(\mathbf{p})) + \frac{1}{4}(r^2 - 1) \right] + y \left[1 + r\mu(\mathbf{p}) \right] = \mu(\mathbf{p})^2 + a_0^2 \tilde{\mathbf{p}}^2 . \quad (\text{C.7})$$

Note that the only way to obtain just one branch is to choose $r=1, b_0=c_0=0$, which corresponds to the Wilson or SW action.

Modulo the slight changes necessary for the ghost-free D234gf action, all dispersion-related results and plots presented throughout this paper can be obtained by analytical or numerical manipulation of eq. (C.5). For completeness we include some more results concerning the small mass expansions of $E(0)/m_c$ and $c(\mathbf{p})$ defined by $E(\mathbf{p})^2 = E(0)^2 + \mathbf{p}^2 c(\mathbf{p})^2$. Here $E(\mathbf{p})$ is the physical branch of the dispersion relation (cf. sect. 4); $c(\mathbf{p})$ is defined only for (sufficiently small) momenta and masses, where there is a well-defined physical branch.

The first few coefficients in the expansions

$$\begin{aligned} \frac{E(0)}{m_c} &= 1 + \sum_{n=1}^{\infty} E_n a_0^n m_c^n \\ c(0) &= 1 + \sum_{n=1}^{\infty} C_n a_0^n m_c^n \end{aligned} \quad (\text{C.8})$$

are presented in tables 1 and 2 for various actions discussed in sect. 3. These results are easily derived by an iterative series expansion of eq. (C.5), using a symbolic manipulation

Action	E_1	E_2	E_3	E_4	E_5	E_6
SW	0	$-\frac{1}{6}$	$\frac{1}{8}$	$-\frac{1}{20}$	0	$\frac{1}{56}$
D34($c_0 = \frac{1}{6}$)	0	0	$\frac{1}{6}$	$\frac{1}{30}$	$\frac{1}{36}$	$\frac{29}{252}$
D234 \equiv W2	0	0	$\frac{1}{18}$	$-\frac{1}{270}$	$\frac{5}{324}$	$\frac{41}{6804}$
D234i($\frac{2}{3}$)	0	0	$\frac{1}{18}$	$-\frac{1}{270}$	$\frac{5}{324}$	$\frac{41}{6804}$
D234($\frac{2}{3}$)	0	0	0	$\frac{1}{30}$	$-\frac{5}{216}$	$\frac{115}{6048}$
D234c($\frac{2}{3}$)	0	0	0	$\frac{1}{30}$	$-\frac{1}{54}$	$\frac{37}{2268}$

Table 1: Expansion coefficients of $E(0)/m_c$, eq. (C.8), for various actions.

Action	C_1	C_2	C_3	C_4	C_5	C_6
SW	0	$-\frac{1}{3}$	$\frac{1}{4}$	$-\frac{7}{180}$	$-\frac{5}{48}$	$\frac{1009}{7560}$
D34($c_0 = \frac{1}{6}$)	0	0	$\frac{1}{3}$	$\frac{1}{10}$	$\frac{1}{12}$	$\frac{22}{63}$
D234 \equiv W2	0	0	$\frac{1}{9}$	$\frac{7}{270}$	$\frac{7}{324}$	$\frac{131}{3402}$
D234i($\frac{2}{3}$)	0	0	$\frac{1}{9}$	$\frac{7}{270}$	$\frac{7}{324}$	$\frac{131}{3402}$
D234($\frac{2}{3}$)	0	0	0	$\frac{1}{10}$	$-\frac{5}{72}$	$\frac{1135}{18144}$
D234c($\frac{2}{3}$)	0	0	0	$\frac{1}{10}$	$-\frac{1}{18}$	$\frac{10}{189}$

Table 2: Expansion coefficients of $c(0)$, eq. (C.8), for various actions.

program. Note that $E(0)/m_c$ and $c(0)$ are independent of the spatial Δ_i^2 terms. For the D34 action we therefore only have to specify c_0 in tables 1 and 2. This also implies that D234i($\frac{2}{3}$) and the isotropic D234 action have the same $E(0)/m_c$ and $c(0)$ expansion coefficients, cf. tables 1 and 2.

For the SW action with $r=1$ the exact mass $E(0)$ is given by the simple formula

$$a_0 E(0) = \ln \left[1 + a_0 m_c \left(1 + \frac{1}{2} a_0 m_c \right) \right]. \quad (\text{C.9})$$

We leave it as an exercise to the reader to derive analytical results for various other special cases.

The dispersion relations plotted throughout this paper apply, strictly speaking, only to the on- but not off-shell improved actions obtained in the third step of our procedure. The reader might wonder what happens to the dispersion relations, when, as instructed in step 3, we undo the change of variable, leading to off-shell improved actions and fields. The off-shell improved propagator in eq. (3.6) implies the dispersion relation

$$\text{Tr}(\Omega(p)\Omega^\star(p)) \text{Tr}(\bar{\Omega}(p)\bar{\Omega}^\star(p)) \left(\tilde{p}^2 + \tilde{m}(p)^2 \right) = 0. \quad (\text{C.10})$$

The factors from undoing field transformation give rise to additional high-lying ghosts, without affecting the other branches of the dispersion relation. Considering, for simplicity, the SW case, where we can choose $\bar{\Omega} = \Omega = 1 - \frac{ra_0}{4}(\nabla - m_c)$, we have

$$\frac{1}{4} \text{Tr}(\Omega(p)\Omega^*(p)) = \frac{1}{4} \text{Tr}(\bar{\Omega}(p)\bar{\Omega}^*(p)) = (1 + \frac{ra_0}{4}m_c)^2 + (\frac{ra_0}{4})^2 \bar{p}^2. \quad (\text{C.11})$$

For zero mass and momentum the extra ghosts are at energies $a_0 E(0) = \text{arcsinh}(4/r)$. If these ghosts are deemed to be not high enough, one can push them up even further by modifying the field transformation operators to read

$$\bar{\Omega} = \Omega = \left[1 - \frac{ra_0}{4n}(\nabla - m_c)\right]^n \quad (\text{C.12})$$

for suitably large integer n . Obviously, similar remarks apply to the D234 actions.

So far we have considered the dispersion relation in the absence of a background gauge field, finding that the coefficients of the scaling errors of our D234 actions are quite small. One might wonder if small coefficients prevail in the presence of a gauge field.

The short answer is yes, because, as for vanishing gauge field, these errors ultimately arise from the truncation of the expansion of the the continuum D_μ in terms of lattice derivatives, which is a series with a finite radius of convergence and rapidly decreasing coefficients.

The longer answer goes as follows. When we discretize the continuum action M_Ω obtained by a change of variable in sect. 3.1, the leading error is of even order, a^2, a^4, \dots , and comes from the truncation of the \not{D} term; the truncation error of \not{D}^2 always comes with an additional factor of a_0 .

The leading error of the general dispersion relation of a D234 action (at least the D234c(r) one, see below) is therefore simply that of the naive improved fermion action $\nabla_c + m_c$. Remembering that $\nabla_{c\mu} = D_\mu - \frac{1}{30}a_\mu^4 D_\mu^5 + \mathcal{O}(a_\mu^6)$ we find the position-space dispersion relation

$$m_c^2 = \sum_\mu D_\mu^2 + \frac{1}{2}\sigma \cdot F - \frac{1}{15} \sum_\mu a_\mu^4 D_\mu^6 - \frac{1}{30} \sum_{\mu\nu} i\sigma_{\mu\nu} [D_\mu, a_\nu^4 D_\nu^5] + \mathcal{O}(a^5). \quad (\text{C.13})$$

The first two terms on the rhs give the continuum dispersion relation, the rest the $\mathcal{O}(a^4)$ errors. Note that the commutator terms vanish in the absence of a gauge field. As promised, the coefficients of the error terms are small.

For the “tuned” D234 actions of sect. 3 there are additional a^3 errors, *e.g.* there is a $-\frac{1}{18}a_0^3\Delta_0^2$ correction to the action for the D234i($\frac{2}{3}$) case. Again, the coefficients of these errors are quite small.

For comparison, here is the dispersion relation of the SW action

$$m_c^2 = \sum_\mu D_\mu^2 + \frac{1}{2}\sigma \cdot F + \frac{1}{3} \sum_\mu a_\mu^2 D_\mu^4 + \frac{1}{6} \sum_{\mu\nu} i\sigma_{\mu\nu} [D_\mu, a_\nu^2 D_\nu^3] + \mathcal{O}(a^3). \quad (\text{C.14})$$

Note that using these expansions it is trivial to calculate the finite-mass corrections to various terms in the action. For example, for the case of a small, constant, *magnetic* background field one finds that the energy of a “D234c particle” at zero velocity is:

$$E^2 = m_c^2 \left(1 + \frac{1}{15} a_0^4 m_c^4\right) - \frac{1}{2} \sigma \cdot F \left(1 + \frac{1}{5} a_0^4 m_c^4\right) + \mathcal{O}(a^5) . \quad (\text{C.15})$$

The first term shows the correction corresponding to the leading coefficient in table 1. The second shows that the correction to the hyper-fine splitting, although not as small as for the mass, is still small; on the one percent level as long as $a_0 m_c < 0.5$. This should be compared with the SW case, where the analogous calculation gives

$$E^2 = m_c^2 \left(1 - \frac{1}{3} a_0^2 m_c^2\right) - \frac{1}{2} \sigma \cdot F \left(1 - \frac{2}{3} a_0^2 m_c^2\right) + \mathcal{O}(a^3) , \quad (\text{C.16})$$

showing that $a_0 m_c$ must be smaller by almost a factor of 4 for the hyper-fine splitting correction to be similarly small.

Appendix D Tadpole Calculus

A large, apparently often dominant, part of the contributions of standard lattice perturbation to a generic quantity are unphysical, in the sense that they are due to tadpole diagrams, which do not exist in standard ways of performing continuum perturbation theory. Lepage and Mackenzie [8] proposed a simple method to improve lattice operators and actions, *i.e.* make them more continuum-like.

At tree level, their prescription amounts to replacing each link field $U_\mu(x)$ in an operator by $U_\mu(x)/u_\mu$, where the number $0 < u_\mu < 1$ should be defined such that, roughly speaking, $1 - u_\mu$ represents the “tadpole part” of the link field. At higher orders, one divides each link by the non-perturbatively measured u_μ , and multiplies by its perturbative expansion. In other words, one reorganizes perturbation theory in a manner that sums up most of the tadpole contributions to all orders, separately from the physical contributions. For isotropic actions all u_μ are the same and a simple, gauge invariant prescription for the u_μ is given by $u_0 = \langle \frac{1}{N} \text{Tr } U_\square \rangle^{1/4}$, where U_\square is the plaquette operator in a theory with $\text{SU}(N)$ gauge fields.

This tadpole improvement prescription is the one used in the literature up to now, and a version for anisotropic lattices was used for exploratory simulations [11, 23]. We are presently also exploring other prescriptions [18].

In the rest of this section we just assume that the u_μ have been chosen by some scheme; their precise numerical values are irrelevant. For our fermionic actions the operators to be tadpole-improved are ∇_μ , Δ_μ and various products and sums thereof. Let $T(A)$ denote the tadpole-improved version of some operator A . Note that T is a linear operator, *i.e.*

$$T(A + \lambda B) = T(A) + \lambda T(B) \quad \text{for any } \lambda \in \mathbb{C} . \quad (\text{D.1})$$

Clearly,

$$\begin{aligned} T(\nabla_\mu) &= \frac{\nabla_\mu}{u_\mu} \\ T(\Delta_\mu) &= \frac{\Delta_\mu}{u_\mu} + \frac{2}{a_\mu^2} \left(\frac{1}{u_\mu^2} - 1 \right). \end{aligned} \quad (\text{D.2})$$

A subtlety arises for sufficiently complicated products of operators. Namely, the improvement prescription requires one to expand out all products in terms of $U_\mu(x)$'s and use $U_\mu(x)U_{-\mu}(x+\mu) \equiv 1$ for “backtracking” products of link fields. This implies that in general

$$T(AB) \neq T(A) T(B). \quad (\text{D.3})$$

In practice it will sometimes be more efficient to expand out a product of operators in terms of link fields — then no subtlety arises. In other cases, however, it will be more convenient or efficient to apply one tadpole-improved operator after another. If we refer to the replacement of a product of ∇_μ 's and Δ_μ 's by the product of their tadpole-improved versions as “naive tadpole improvement”, the simplest strategy in such a case is to first apply naive tadpole improvement and then correct the error by using the following formulas:

$$\begin{aligned} T(\nabla_\mu \Delta_\mu) &= T(\nabla_\mu) T(\Delta_\mu) \\ T(\nabla_\mu^2) &= T(\nabla_\mu)^2 + \delta_\mu \\ T(\Delta_\mu^2) &= T(\Delta_\mu)^2 - \frac{4}{a_\mu^2} \delta_\mu \\ T(\nabla_\mu^3) &= T(\nabla_\mu)^3 + \frac{3}{2} \delta_\mu T(\nabla_\mu) \\ T(\Delta_\mu^3) &= T(\Delta_\mu)^3 - \frac{6}{a_\mu^2} \delta_\mu T(\Delta_\mu) + \frac{12}{a_\mu^4} \delta_\mu \end{aligned} \quad (\text{D.4})$$

where

$$\delta_\mu \equiv \frac{1}{2a_\mu^2} \left(\frac{1}{u_\mu^2} - 1 \right). \quad (\text{D.5})$$

We also remind the reader of the identities (B.7) and (B.8). Easy corollaries of eqs. (D.4) are

$$\begin{aligned} T(\nabla^2) &= T(\nabla)^2 + \sum_\mu \delta_\mu \\ T(\nabla^3) &= T(\nabla)^3 + \left(\sum_\mu \delta_\mu \right) T(\nabla) + \frac{1}{2} \sum_\mu \delta_\mu T(\gamma_\mu \nabla_\mu) \end{aligned} \quad (\text{D.6})$$

which are useful for the tadpole improvement of the field redefinition operators Ω and $\bar{\Omega}$.

References

- [1] S. Gupta, A. Irbäck, F. Karsch and B. Petersson, Phys. Lett. **B242** (1990) 437.
- [2] M. Alford, W. Dimm, G.P. Lepage, G. Hockney, P.B. Mackenzie, Phys. Lett. **B361** (1995) 87.
- [3] K. Symanzik, in: *Mathematical Problems in Theoretical Physics*, R. Schrader *et al.* (eds.), LNP 153, Springer, Berlin, 1982; in: *Recent Developments in Gauge Theories*, G. 't Hooft (ed.), Plenum Press, New York, 1980; Nucl. Phys. **B226** (1983) 187, 205.
- [4] M. Lüscher and P. Weisz, Comm. Math. Phys. **97** (1985) 59, (E) **98** (1985) 433.
- [5] M. Lüscher and P. Weisz, Phys. Lett. **158B** (1985) 250.
- [6] B. Sheikoleslami and R. Wohlert, Nucl. Phys. **B259** (1985) 609.
- [7] M. Lüscher, S. Sint, R. Sommer, P. Weisz, H. Wittig and U. Wolff, hep-lat/9608049 and references therein.
- [8] G.P. Lepage and P.B. Mackenzie, Phys. Rev. **D48** (1993) 2250.
- [9] C.T.H. Davies, K. Hornbostel, G.P. Lepage, A.J. Lidsey, J. Shigemitsu and J. Sloan, Phys. Rev. **D52** (1995) 6519.
- [10] M. Alford, T.R. Klassen, and G.P. Lepage, Nucl. Phys. **B** (Proc. Suppl), **47** (1996) 370.
- [11] M. Alford, T.R. Klassen, and G.P. Lepage, hep-lat/9608113.
- [12] S. Collins, R. Edwards, U. Heller and J. Sloan, hep-lat/9608021.
- [13] A. Borici and Ph. de Forcrand, hep-lat/9608105.
- [14] W. Bock, hep-lat/9608103.
- [15] B. Beinlich, F. Karsch, and A. Peikert, hep-lat/9608141.
- [16] K.G. Wilson, in *New Phenomena in Subnuclear Physics*, Part A, A. Zichichi (ed.), Plenum Press, New York, p. 69, 1975.
- [17] F. Karsch, Nucl. Phys. **B205** (1982) 285; G. Burgers, F. Karsch, A. Nakamura, and I.O. Stamatescu, Nucl. Phys. **B304** (1988) 587.
- [18] M. Alford, T.R. Klassen, G.P. Lepage, C. Morningstar, M. Peardon, and H. Trottier, *QCD on Anisotropic Lattices*, to appear.
- [19] G. Heatlie, C.T. Sachrajda, G. Martinelli, C. Pittori and G.C. Rossi, Nucl. Phys. **B352** (1991) 266.

- [20] G.P. Lepage, L. Magnea, C. Nakhleh, U. Magnea and K. Hornbostel, Phys. Rev. **D43** (1992) 4052.
- [21] A.X. El-Khadra, A.S. Kronfeld and P.B. Mackenzie, [hep-lat/9604004](#).
- [22] M. Alford, T.R. Klassen and G.P. Lepage, work in progress.
- [23] C. Morningstar and M. Peardon, [hep-lat/9608050](#), [hep-lat/9608019](#).
- [24] G.P. Lepage, Nucl. Phys. **B** (Proc. Suppl), **47** (1996) 3.
- [25] G.P. Lepage, Schladming Winter School lectures, [hep-lat/9607076](#).
- [26] H.B. Nielsen and M. Ninomiya, Nucl. Phys. **B185** (1981) 20; **B193** (1981) 173.
- [27] R. Narayanan and H. Neuberger, Nucl. Phys. **B443** (1995) 305.
- [28] L.H. Karsten and J. Smit, Nucl. Phys. **B183** (1981) 103.
- [29] H.W. Hamber and C.M. Wu, Phys. Lett. **133B** (1983) 351; T. Eguchi and N. Kawamoto, Nucl. Phys. **B237** (1984) 609.
- [30] H.R. Fiebig and R.M. Woloshyn, [hep-lat/9603001](#); R. Lewis and R.M. Woloshyn, [hep-lat/9610027](#); F. Lee and D. Leinweber, [hep-lat/9606005](#).
- [31] We would like to thank J. Sloan and R. Edwards for discussions on this.